African trypanosomes evade immune clearance by O-glycosylation of the VSG surface coat

Pinger, Jason; Nešić, Dragana; Ali, Liaqat; Aresta-Branco, Francisco; Lilic, Mirjana; Chowdhury, Shanin

Published in:
Nature Microbiology

DOI:
10.1038/s41564-018-0187-6

Publication date:
2018

Document Version
Peer reviewed version

Supplementary Information

African trypanosomes evade immune clearance by *O*-glycosylation of the VSG surface coat

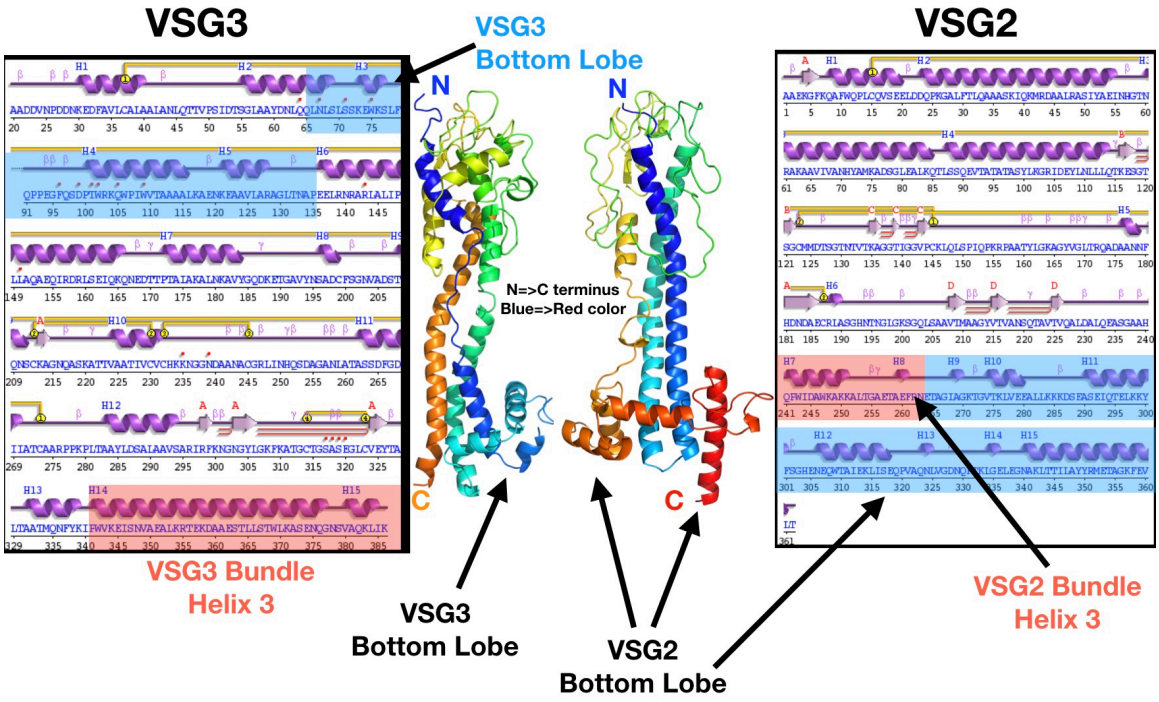
Jason Pinger[1*], Dragana Nešić[2*†], Liaqat Ali[3*], Francisco Aresta-Branco[4,5], Mirjana Lilic[2‡], Shanin Chowdhury[1], Hee-Sook Kim[1], Joseph Verdi[6], Jayne Raper[6], Michael A. J. Ferguson[3§], F. Nina Papavasiliou[4§], C. Erec Stebbins[5§]

[1]The Rockefeller University, Laboratory of Lymphocyte Biology, New York, New York, USA [2]The Rockefeller University, Laboratory of Structural Microbiology, New York, New York, USA. [3]Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dundee, UK. [4]Division of Immune Diversity, German Cancer Research Center, Heidelberg, Germany. [5]Division of Structural Biology of Infection and Immunity, German Cancer Research Center, Heidelberg, Germany. [6]Department of Biological Sciences, Hunter College, City University of New York, New York, USA

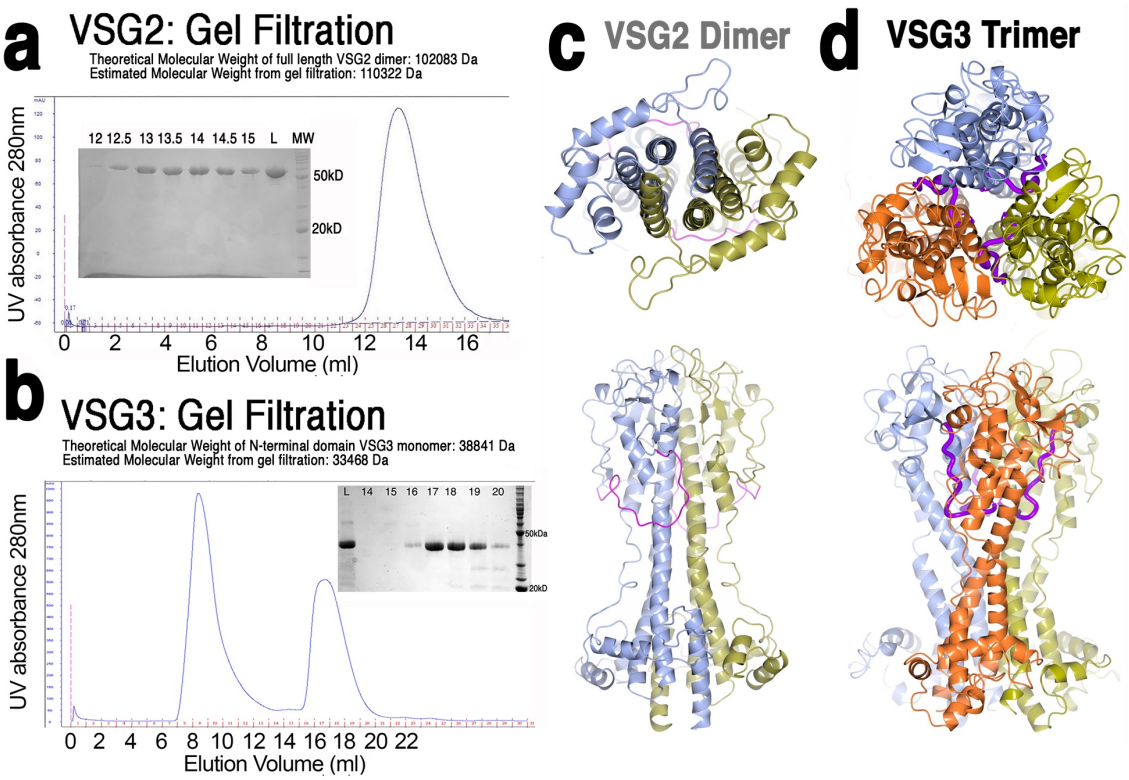
[*]These authors contributed equally to this work.

[†]Current address: The Rockefeller University, Laboratory of Blood and Vascular Biology, New York, New York, United States of America. [‡]Current address: The Rockefeller University, Laboratory of Molecular Biophysics, New York, New York, United States of America.

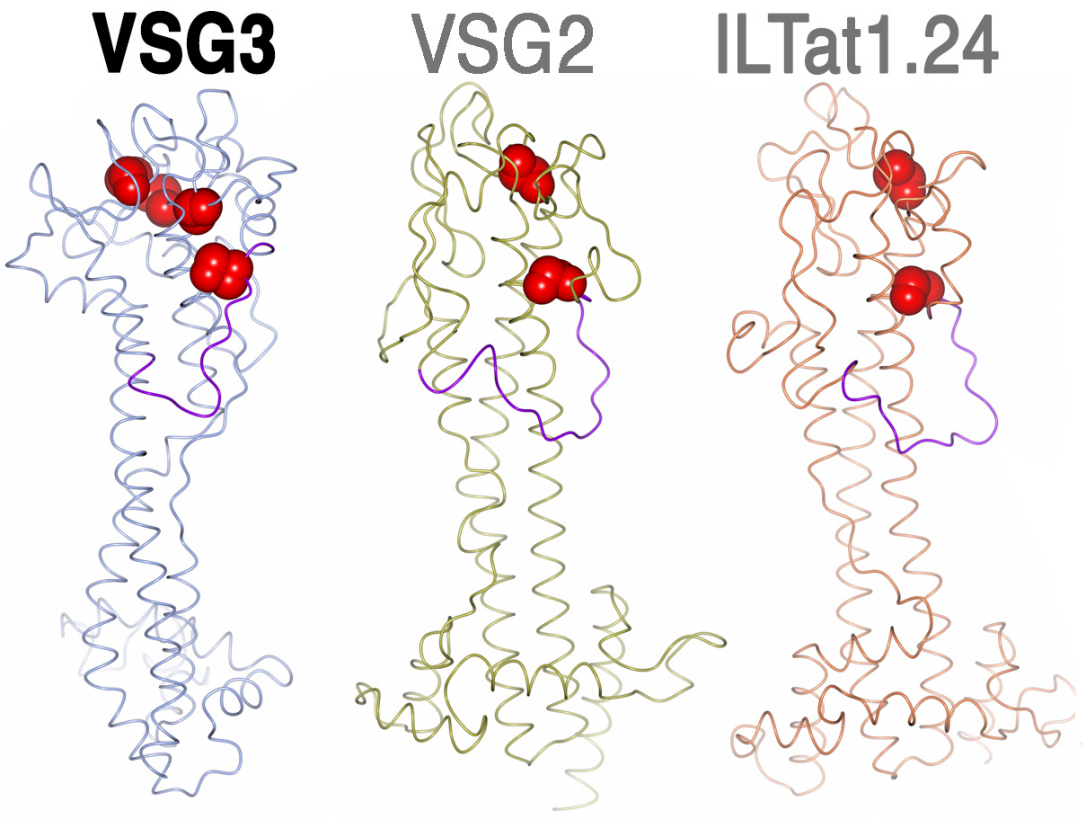
[§]Correspondence to: e.stebbins@dkfz-heidelberg.de (structural biology), n.papavasiliou@dkfz-heidelberg.de (trypanosome strains and infection assays), and m.a.j.ferguson@dundee.ac.uk (carbohydrate biochemistry)



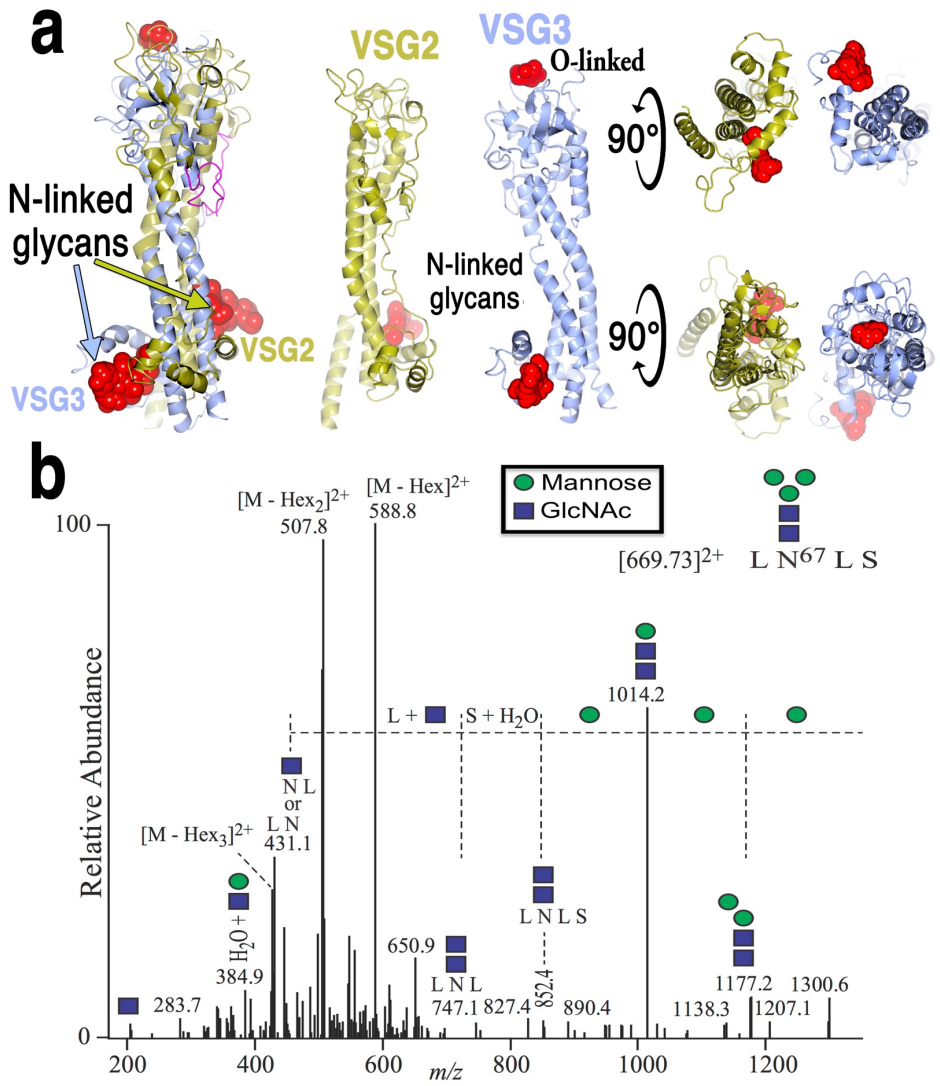
Supplementary Fig. 1: Topological Divergence Lower Lobe Placement and Structure. Shown are the sequence and secondary structural elements of the monomers of VSG3 (left) and VSG2 (right) with annotations. Between the sequences are ribbon diagrams of the respective monomeric structures colored in a gradient from blue to red from the N to the C-terminus. Of note is the divergent connectivity of the lower-lobe subdomain in each polypeptide sequence (the subdomain modified with an N-linked carbohydrate in both and which is proximal to the membrane-anchored C-terminal domain). In VSG3, the smaller bottom lobe occurs near the N-terminus of the primary sequence immediately following the first helix of the three-helix bundle (stalk helix). In VSG2, the bottom lobe is the final stretch of polypeptide in the protein, immediately following the third and final helix of the bundle. In both sequences, the bottom lobe regions are shown with blue backgrounds and the third bundle helices are shown with red backgrounds. In the ribbon diagrams, the N-terminal to C-terminal color gradient from blue to red makes the sequence positioning of the bottom lobe clearer: in VSG2 the bottom lobe is orange-red as it occurs at the end of the polypeptide sequence, whereas in VSG3 it is blue-cyan as the bottom lobe occurs near the beginning of the polypeptide sequence.



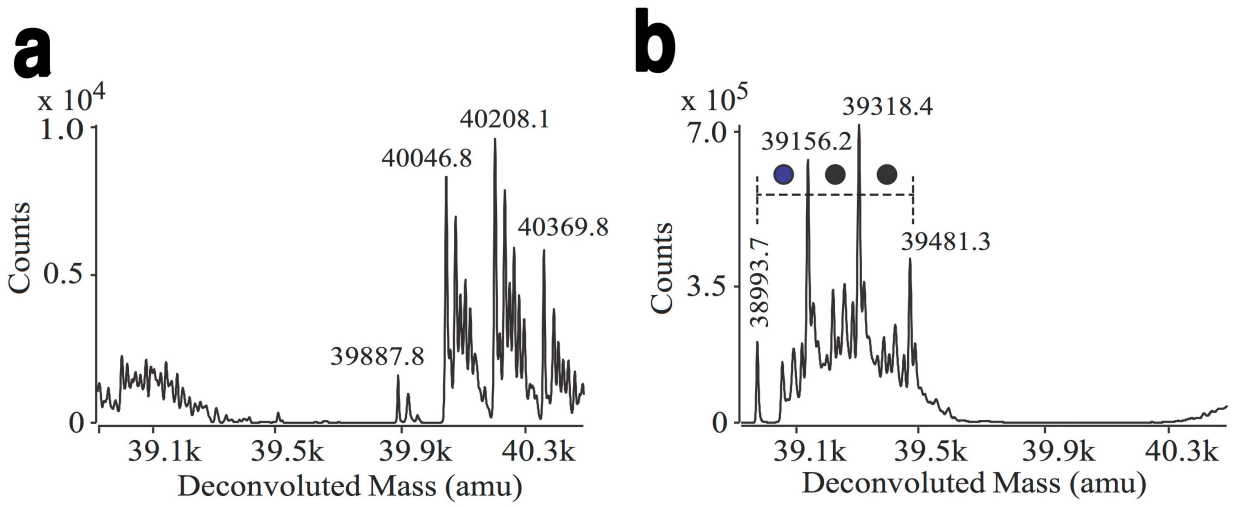
Supplementary Fig. 2: Oligomerization differences between VSGs.
(a) Gel filtration chromatography of full length VSG2 (24ml Superdex 200 column, GE Healthcare). Pasted beside the chromatogram is a coomassie stained SDS-PAGE gel of the corresponding peak fractions shown and labeled by elution volume. The calculated molecular weight from the elution volumes (based on a calibration curve produced from multiple protein standards) indicates that the peak corresponds to a mass of approximately 110kDa, near the predicted polypeptide molecular weight of a VSG2 dimer of 102kDa. On the gel, L=Load (the material applied to the column), MW=peptide molecular weight standards, and the numbers above each lane refer to the elution volume. **(b)** Gel filtration chromatography of the N-terminal domain of VSG3 (24ml Superdex 200 column, GE Healthcare) with coomassie stained SDS-PAGE gel of peak fractions shown and labeled by elution volume. The peak corresponds to a mass of approximately 33.5kDa, near the predicted molecular weight of the N-terminal domain of the VSG3 monomer of 38.8kDa. The large peak beginning at around 7ml is the "void volume" of the column, and corresponds to soluble aggregates of millions of Daltons in size. On the gel, L=Load (the material applied to the column) and the numbers above each lane refer to the elution volume. Gel filtration experiments were performed repeatedly (many more than three times) during the course of producing protein for crystallization trials for both VSG2 and VSG3 with similar results to what is shown here.
Comparison of the VSG2 dimer **(c)** and the VSG3 crystallographic trimer **(d)**. Individual monomers are colored differently (blue, yellow, and orange). Two views, a "top" and "side", are shown. The conserved loop in the upper lobe is colored purple, showing that its orientation is inward facing for the VSG3 trimer, and outward facing for the VSG2 dimer. While VSG2 and ILTat1.24 have formed non-crystallographic dimers, VSG3 was found to oligomerize only through a three-fold crystallographic axis. This trimer differs markedly from the dimers of previous VSGs, with the faces of VSG3 that correspond to the dimerization interfaces of VSG2 and ILTat1.24 oriented outward and away from the other elements in the trimer. The trimer observed in the crystals is composed of individual monomer interactions each with nearly 1500 A^2 of buried surface area. The biological significance of this crystallographic trimer remains to be examined.



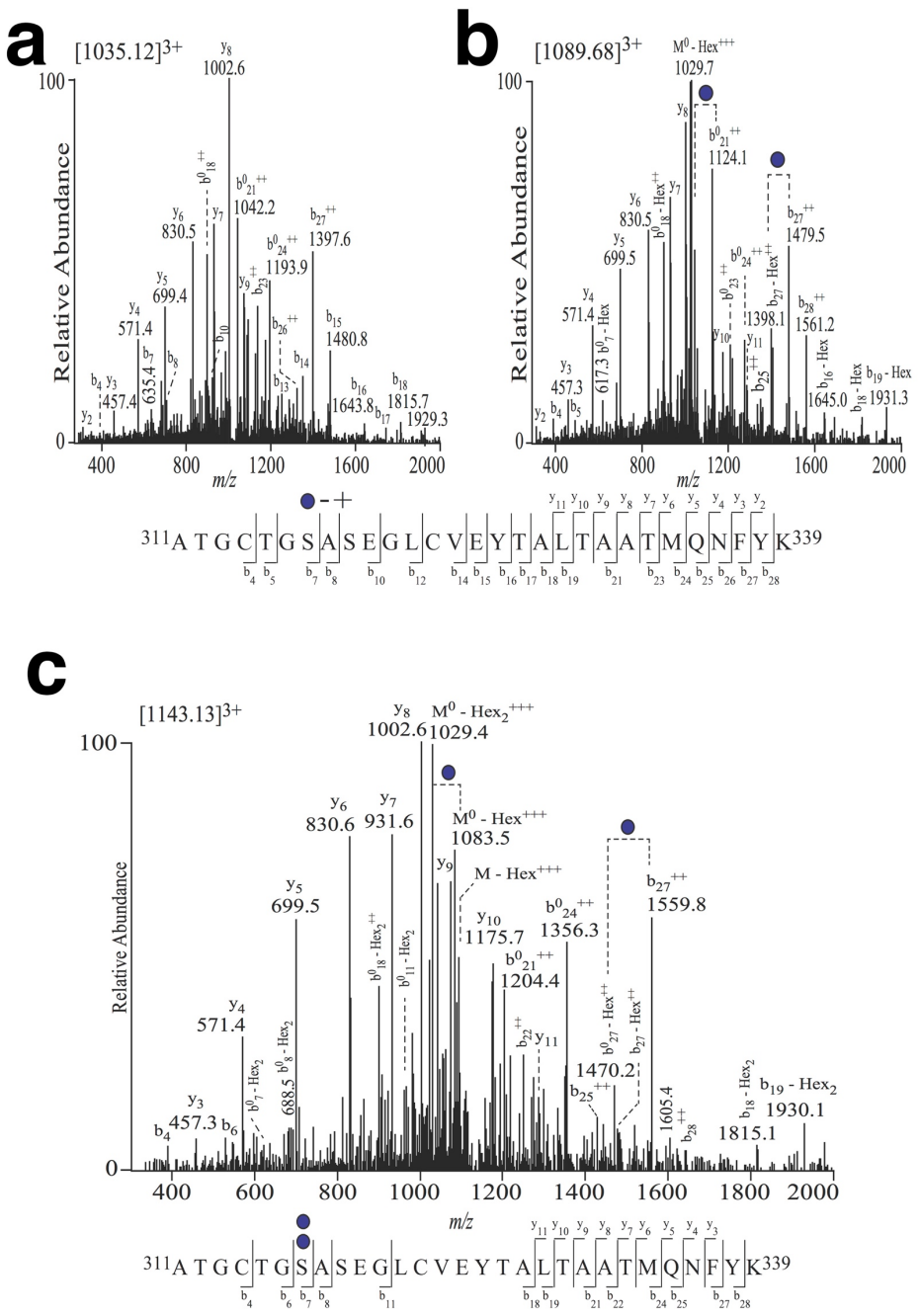
Supplementary Fig. 3: Comparisons of VSG2 and VSG3 Disulfide Bonds.
The cysteine disulfide bonds (red, space-filling atoms) are shown in the context of a ribbon diagram of VSG3, VSG2, and ILTat1.24. VSG3 has 9 cysteines, 8 of which are in S-S bonds, one is free: C199 which is buried in a hydrophobic pocket with Y184, Y194, and A225.



Supplementary Fig. 4: MS/MS spectrum of the N67 Pronase glycopeptide.
(a) Similar to 2a from main text, shown here as visual context for discussion of lower lobe N-glycans in VSG2 and VSG3 (gold and blue ribbon structures indicate VSG2 and VSG3, respectively). VSGs generally contain 1-3 occupied N-glycosylation sites, with one of these often occurring at the end of the N-terminal domain (in the lower lobe of the structure). However, there is structural diversity in this lower lobe, even in the more structurally similar VSG2 and ILTat1.24; for example, the space occupied by a short peptide helix in ILTat1.24 (not shown) is instead occupied by the core of a family of small paucimannose/complex N-glycans in VSG2 (lower lobe VSG2 glycan shown here). A similar small paucimannose N-glycan was identified at N67 for VSG3 but on the opposite face of the structurally divergent lower lobe (the structures are taken from the alignment in 2a, separated, but their relative orientation in the alignment maintained). The structure is predominantly Man3GlcNAc2, as deduced by ES-MS/MS analysis of a Pronase glycopeptide fraction of VSG3 (shown in (b)) and by analysis of the intact glycoprotein before and after de-N-glycosylation with PNGaseF. VSG3 also has an unoccupied potential N-glycosylation sequon at N86. **(b)** LC-MS/MS analysis of the Pronase glycopeptide fraction revealed a doubly-charged ion at *m/z* 669.73 consistent with it being the $[M+2H]^{2+}$ ion for $Hex_3HexNAc_2$-(LNLS), which includes the N67 N-glycosylation sequon. The MS/MS spectrum shown here further confirmed this identity. This experiment was performed once.

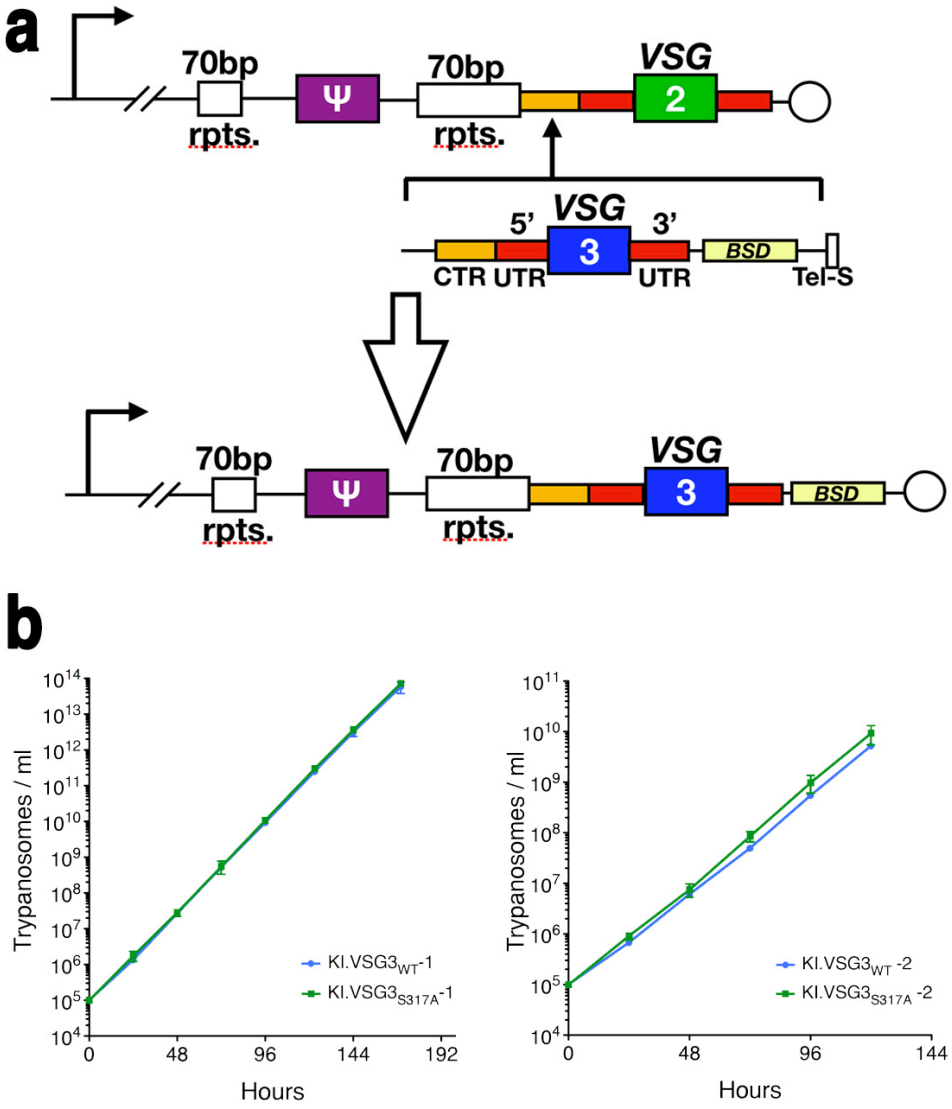


Supplementary Fig. 5: Confirmation of *O*-linked glycan on VSG3
The deconvolved electrospray mass spectra of the purified native VSG3 domain used for crystallization before **(a)** and after **(b)** de-N-glycosylation with PNGaseF. The mass differences of approximately 162 Da between the VSG3 species in both cases indicate *O*-glycosyl heterogeneity. The lowest mass in panel (b), 38993.7 Da, agrees with the theoretical mass for residues A20-G388 of VSG3 allowing for the conversion of N67 to D67 by PNGaseF and for four disulfide bonds. The experiment shown in (a) was performed three times with similar results, and the experiment shown in (b) was performed twice with similar results.

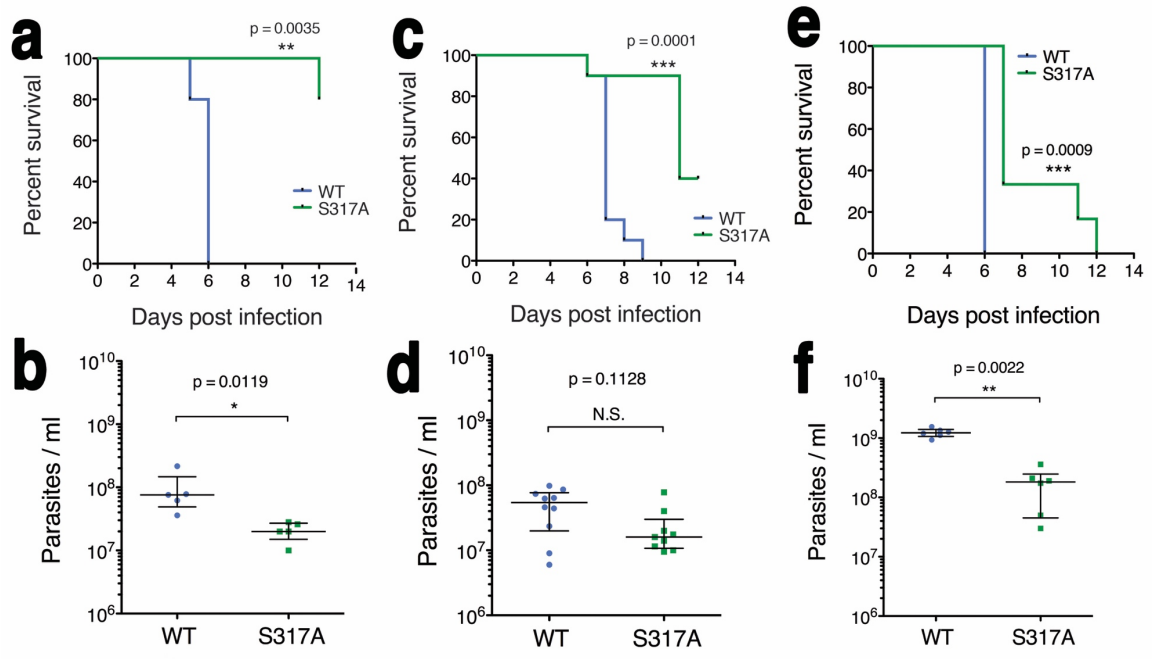


Supplementary Fig. 6: MS/MS of heterogeneous *O*-linked VSG3 surface glycans.

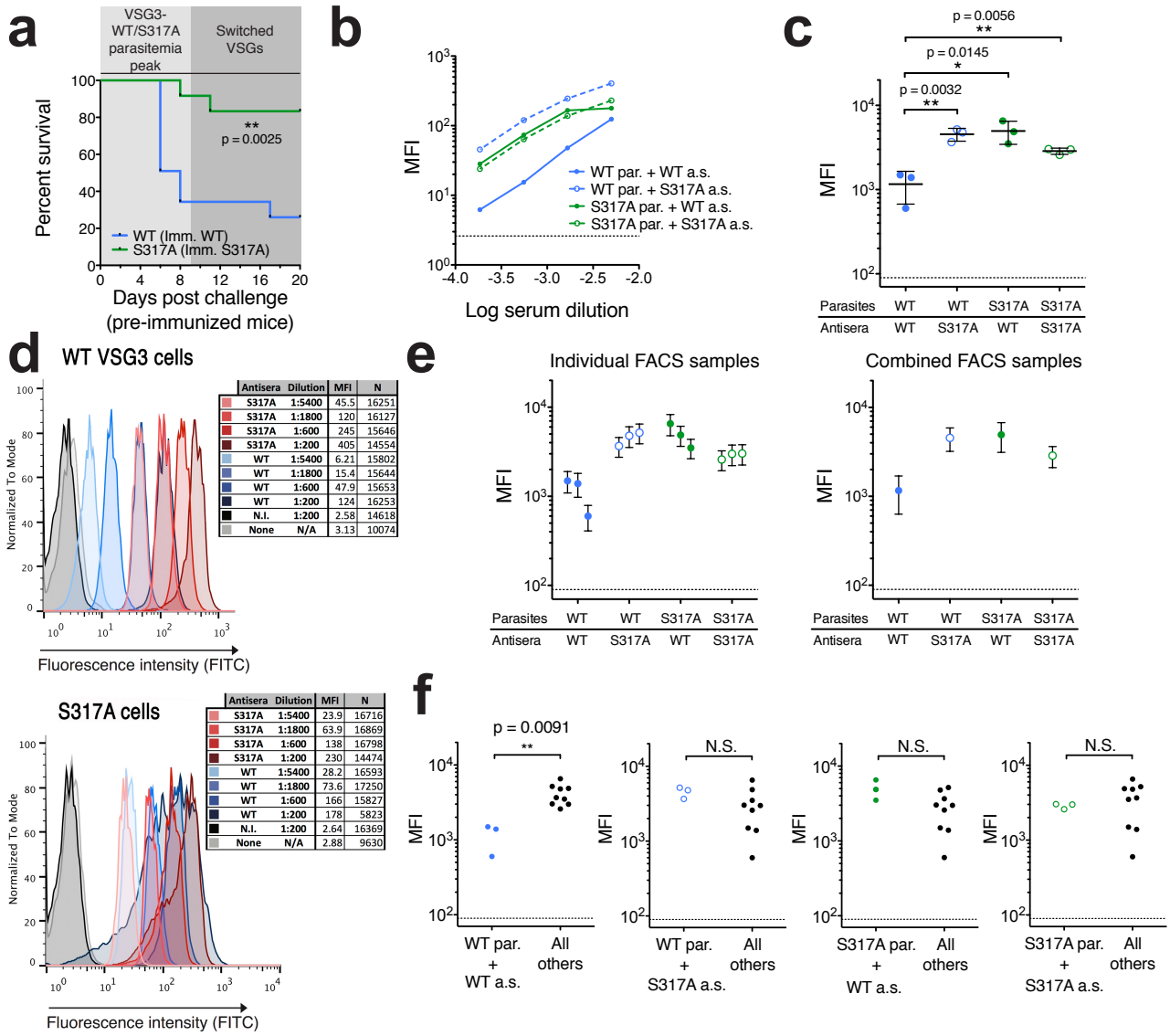
(a) - **(c)** are the product ion mass spectra of $[M+3H]^{3+}$ ions corresponding to the VSG3 peptide A311-K339 containing **(a)** zero, **(b)** one, and **(c)** two hexose residues. For (b), the diagnostic b7-Hex ion at *m/z* 617.3 suggests that S7 in the peptide was converted to dehydro-alanine through β-elimination, defining the *O*-glycosylation site in the original VSG3 sequence as S317. For (c), the b7 and b8 ions suggest that both hexoses are attached to S7 (S317 in the full VSG3 sequence) and not to any other hydroxyl-amino acids in the sequence. Each of these experiments were performed twice with similar results.



Supplementary Fig. 7: Creation of isogenic VSG3$_{WT}$ and VSG3$_{S317A}$ expressing strains. **(a)** Schematic description of the generation of isogenic VSG3$_{WT}$ or VSG3$_{S317A}$ expressing strains. The active VSG expression site of the parental clones [*Lister427* Bloodstream-form Expression site 1 (BES1)] is depicted, with the 70bp repeats and telomere (circle) abutting the co-transposed region (CTR, orange) and UTRs of *VSG2*. Replacement of *VSG2* by *VSG3* (wild type or mutant) was achieved by homologous recombination using a template containing the entire *VSG3* gene (including UTRs), together with a Blasticidin (BSD) resistance cassette for selection purposes, flanked by the BES1 *VSG2* CTR and a telomere seed (Tel-S). The resulting clones mimic a naturally occurring VSG switch event, except with the addition of the BSD cassette. **(b)** Trypanosome clones expressing VSG3$_{WT}$ or VSG3$_{S317A}$ mutant coats grow at comparable rates *in vitro* (left: data for clones KI.VSG3$_{WT}$-1 and KI.VSG3$_{S317A}$-1; right: data for clones KI.VSG3$_{WT}$-2 and KI.VSG3$_{S317A}$-2). Cultures were counted and diluted daily and cell densities were maintained between 10^5 and 2 x 10^6 cells /ml. Data show overall growth, calculated by multiplying the concentration each day by the dilution factors of previous days Data are plotted as mean ± SD of 4 independent growth curves for clones KI.VSG3$_{WT}$-1 and KI.VSG3$_{S317A}$-1, and 3 independent curves for clones KI.VSG3$_{WT}$-2 and KI.VSG3$_{S317A}$-2.



Supplementary Fig. 8: Mouse infection assays.
(a) and **(b)** show data from infection of C57BL/6 (inbred) mice with trypanosome clones KI.VSG3$_{WT}$-1 and KI.VSG3$_{S317A}$-1 (n = 5 mice/group), while **(c)** and **(d)** show data from infection of C57BL/6 (inbred) mice with two additional trypanosome clones, KI.VSG3$_{WT}$-2 and KI.VSG3$_{S317A}$-2 (n = 10 mice/group). KI.VSG3$_{WT}$-2 and KI.VSG3$_{S317A}$-2 were generated on a GPIPLC$^{-/-}$ background, which was necessary for the immunization experiments described in Fig. 4c and Supplementary Fig. 9a; this background does not affect parasite virulence or the course of parasitemia[1] (see Methods). Mouse survival is shown in panels (a) and (c), and parasitemia at day 5 post-infection is shown in panels (b) and (d). The aggregate of this data is presented in Figure 4a and 4b. **(e)** and **(f)** show data from infection of CD-1 (outbred) mice with trypanosome clones KI.VSG3$_{WT}$-1 and KI.VSG3$_{S317A}$-1 (n = 6 mice/group). Data are again presented as (e) mouse survival and (f) parasitemia at day 5 post-infection. Statistical analysis was performed using the two-sided log-rank (Mantel-Cox) test for (a), (c), and (e); and by two-sided Mann-Whitney test for (b) (d) and (f) (bars represent median and interquartile range).



Supplementary Fig. 9: Effect of VSG3 *O*-glycosylation on the functionality of host antibodies.
(a) Survival of pre-immunized mice following challenge with live parasites (n = 12 mice/group). Mice were inoculated with UV-irradiated trypanosomes expressing $VSG3_{WT}$ or $VSG3_{S317A}$ (Imm. WT or S317A, respectively) on days -8 and -5 relative to challenge. Challenge injection contained live parasites expressing the same VSG to which the mice were immunized. Shaded areas represent sequential "peaks" of parasitemia. Statistical analysis was performed using the two-sided log-rank (Mantel-Cox) test. **(b-f)** Comparison of $VSG3_{WT}$ or $VSG3_{S317A}$-elicited IgM antisera binding to live $VSG3_{WT}$ or $VSG3_{S317A}$-expressing parasites. Binding was measured by FACS analysis using dilutions of primary antisera, visualized after counterstaining with FITC-labeled anti-mouse-IgM, and reported as mean fluorescence intensity (MFI). **(b)** and **(c)** are the same as Fig. 4d and e from the main text, but include additional cross-reactivity data of each antiserum sample binding parasites expressing the opposing VSG as indicated. **(b)** Data are shown for dilutions of one antiserum sample per VSG type. This experiment was performed twice with similar results. **(c)** Similar to (b) but with a single dilution (1:2000) of distinct antisera, elicited by infection of multiple mice (n = 3 independent antisera/VSG type). Statistical values were calculated by two-sided *t*-tests comparing the sets of three sample MFIs, assuming normal distribution (bars show mean and SD), and similarity of variances was confirmed by F test. **(d)** These plots represent "raw" data used to derive Fig. 4c (dilutions of one antiserum sample per VSG type). Legends indicate antiserum type and dilution, sample MFI, and number of cells analyzed (N). "Antisera - None" indicates cells incubated without primary

antisera or secondary antibody; "N.I." indicates cells incubated with primary antiserum isolated from a naïve, Non-Immunized mouse. These experiments were performed twice with similar results.
(e) (Left) The mean (MFI) and standard deviation (SD; error bars) of the FACS data shown in Fig. 4e and panel (b) of the present figure (three independent antiserum samples per VSG type; one dilution (1:2000) of each). (Right) MFI and SD of the same FACS data combined by sample type (i.e. independent antiserum samples grouped). These data are also shown numerically in Supplementary Table 2 below, which includes the number of cells analyzed (N) in each sample/group. Because FACS sample Ns are so large, a two-tailed *t*-test or non-parametric (Mann-Whitney) test of each individual sample compared to each other sample (and each group compared to each other group) yields a p value < 0.001 for every comparison. These p values signify real variation between samples, but do not inform the magnitude or relevance of these variations. Two comparisons offer a more meaningful interpretation of the data: (1) The mean of each individual WT parasites + WT antiserum sample is at least one SD below the mean of each individual sample from all other groups, and (2) The mean of the combined WT parasites + WT antisera group is two SDs below the mean of each other group (calculated using the larger of the two SDs in each comparison). Therefore, the differences are not only statistically significant (p values), but the WT cells + WT antisera combination shows meaningfully lower IgM binding than all other combinations, based on the magnitude of the differences in sample means compared to the degree of variation within the data. **(f)** Data in main Fig 4e and Supplementary Figure 9e examined differently by regrouping sample MFIs and performing a nonparametric, two-tailed Mann-Whitney test of each group of 3 compared to the other 9 samples combined (as labeled; par. = parasites, a.s. = antisera). Through this analysis, the WT cells + WT antisera combination is the *only* group that is significantly different from all the others.

Supplementary Table 1

VSG3 Crystallography	
Data Collection	
Space group	*I*23
Cell dimensions	
a, b, *c* (Å)	129.8 129.8 129.8
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	91.82-1.41 (1.46-1.41)
Completeness (%)	97.9(73.9)
Redundancy	4.1(1.6)
CC(1/2)	1.0 (0.41)
$I/\sigma I$	19.1 (0.6)
R_{sym}	4.1 (83.8)
Refinement	
Resolution (Å)	91.82-1.41
No. reflections	64954
R_{work} (%)	13.0
R_{free} (%)	16.9
Average B-factor	23.2
R.m.s deviations	
Bond lengths (Å)	0.017
Bond angles (°)	1.758
No. non-hydrogen atoms	3188
Protein	2652
Ligands	61
Solvent	475
Ramachandran Plot	
Favored (%)	98
Allowed (%)	2
Outlier (%)	0

Highest-resolution shell statistics are in parentheses.

		Sample 1			Sample 2			Sample 3		
Parasites	**Antisera**	**MFI**	**SD**	**N**	**MFI**	**SD**	**N**	**MFI**	**SD**	**N**
WT	**WT**	1493	404	17763	1391	416	17829	600	191	17832
WT	**S317A**	3663	916	17729	4783	1250	17323	5181	1298	17533
S317A	**WT**	6533	1753	16493	4874	1234	17579	3500	859	18039
S317A	**S317A**	2588	652	18121	2989	779	17742	3022	776	17846

		Combined		
Parasites	**Antisera**	**MFI**	**SD**	**N**
WT	**WT**	1162	532	53424
WT	**S317A**	4542	1333	52585
S317A	**WT**	4928	1808	52111
S317A	**S317A**	2867	764	53709

Supplementary Table 2. Effect of VSG3 *O*-glycosylation on host antibody binding – FACS values. The mean (MFI), standard deviation (SD), and number of cells analyzed (N) for each sample/group of the FACS data shown in Fig. 4e and Supplementary Fig. 9c, e, and f (three independent antiserum samples per VSG type; one dilution (1:2000) of each).